



PROJETO DE PESQUISA - INICIAÇÃO CIENTÍFICA

ESTÁGIO CIENTÍFICO E TECNOLÓGICO II - EE016

Predição de Séries Temporais Baseada em Redes Neurais Artificiais

Submetido à
Faculdade de Engenharia Elétrica e de Computação (FEEC)

Departamento de Engenharia de Computação e Automação Industrial (DCA)
Faculdade de Engenharia Elétrica e de Computação (FEEC)
Universidade Estadual de Campinas (UNICAMP)
CEP 13083-852, Campinas - SP

Aluno: João Pedro de Oliveira Pagnan
Orientador: Prof. Levy Boccato

Campinas, 28 de julho de 2021

Resumo

Neste trabalho, consideraremos o problema de predição de séries temporais baseada em redes neurais artificiais, que são estruturas bio-inspiradas de processamento dotadas de uma enorme capacidade de elaborar mapeamentos não-lineares. Serão consideradas a clássica rede MLP e algumas propostas de redes recorrentes, como as LSTM e as redes com estados de eco. Diferentes estratégias de aprendizado serão testadas em cenários de predição com características distintas, sendo o objetivo principal desse estudo o estabelecimento de bases para uma análise comparativa das metodologias, análise esta que tem um valor acadêmico e formativo significativo.

1 Introdução e Justificativa

A tarefa de prever valores futuros de uma série temporal a partir de valores presentes e passados é uma das aplicações mais instigantes em tratamento da informação [1]. Tal problema se reveste de uma importância decisiva numa gama de aplicações práticas que envolvem áreas tão distintas quanto econometria e automação industrial, ou meteorologia e biologia populacional. De certa forma, tal interesse prático se liga a dois anseios fundamentais: (1) o desejo de antecipar padrões de comportamento com a finalidade de estabelecer metodologias racionais de decisão, e (2) a busca por construir modelos adequados de fenômenos diversos, de modo a possibilitar uma maior compreensão acerca do comportamento dinâmico subjacente aos próprios dados, o que reveste a tarefa de uma importância teórica e prática ainda maior.

Tipicamente, o problema de predição é abordado por meio de um receituário com três etapas fundamentais (que não precisam necessariamente seguir uma ordem rígida) [2]: escolha de uma estrutura de filtragem que fará o papel de modelo de predição ou *preditor*, adoção de um critério que expresse a qualidade da predição efetuada pelo dispositivo escolhido e emprego de um método de otimização que, tendo em vista o binômio estrutura/critério adotado, forneça um conjunto ótimo de parâmetros livres a serem usados pelo preditor. A segunda dessas etapas é, em geral ¹, definida a partir do uso de um critério de erro quadrático médio (EQM) de predição, que origina a seguinte função custo [4]:

$$J_{EQM} = E\{(x(n + P) - y(n))^2\}, \quad (1)$$

sendo $E\{\cdot\}$ o operador *esperança matemática*, $y(n)$ a saída do preditor no instante n e $x(n + P)$ o valor da série analisada no instante $n + P$. Nesse caso, assume-se que a predição se dá P passos à frente e que o mapeamento expresso por $y(n)$ depende de $x(n)$ e, possivelmente, de valores passados.

Por outro lado, a definição da estrutura e do algoritmo de otimização adotados depende de uma reflexão maior acerca da natureza do problema em mãos. Embora seja comum (e, até certo ponto, bastante útil), inicialmente recorrer ao clássico ferramental linear, como é o caso do modelo autorregressivo (AR) [4], é preciso, em muitos casos

¹Devem ser mencionados, não obstante, esforços bastante meritórios e promissores no sentido de empregar grandezas derivadas da teoria de informação no processo de elaboração de critérios de filtragem [3].

práticos, adotar estruturas de filtragem de cunho *fortemente não-linear*.

Nesse caso, surgem como candidatas interessantes ao papel de estrutura de predição as *redes neurais artificiais* [4], dispositivos bio-inspirados de processamento cujo *modus operandi* se baseia em determinadas características do sistema nervoso. Consideramos particularmente relevantes as redes neurais do tipo MLP (do inglês *multi-layer perceptron*) [5]. Em casos em que há a marca da necessidade de introdução de um caráter dinâmico, também devem ser levadas em conta as *redes recorrentes* [6] (RNNs, do inglês *recurrent neural networks*), que podem corresponder a versões das redes supracitadas em que há a presença de laços de realimentação.

Neste contexto, um modelo recorrente de destacada importância na literatura é a chamada *long short-term memory* (LSTM) [7], a qual apresenta um potencial de recordar valores dentro de sua estrutura relacionados a horizontes de tempo arbitrários. Em outras palavras, ela é capaz de memorizar informações tanto de curto quanto de longo prazo do histórico de eventos relacionados ao problema. Por sua vez, a GRU (*Gated Recurrent Unit*) também explora a noção de porta (*gate*) na definição do comportamento da unidade de processamento, mas de uma forma simplificada em relação à LSTM, obtendo também resultados promissores em algumas aplicações [8].

Outra perspectiva promissora com RNNs refere-se às redes neurais com estados de eco (ESNs, do inglês *echo state networks*). Estas estruturas combinam a capacidade de processamento dinâmico de RNNs com a simplicidade de treinamento típica de modelos lineares [9].

Tendo em vista a diversidade de opções trazidas pela ideia de usar um *framework* baseado em redes neurais para a predição de séries, parece-nos bastante interessante conduzir um estudo em que as duas opções expostas sejam comparadas no âmbito de um conjunto cenários relevantes: esse é o objetivo central deste trabalho. O estudo, além de ter um significativo valor acadêmico, também propiciará ao candidato a oportunidade de adquirir uma formação sólida na área de redes neurais, bem como na linha da pesquisa de séries temporais como um todo.

2 Objetivos

O primeiro objetivo deste trabalho é a realização de um estudo teórico envolvendo sinais e sistemas, elementos de processos estocásticos, filtragem adaptativa supervisionada, redes neurais, aprendizado de máquina e elementos de modelamento / predição de séries temporais. Em particular, pretendemos cobrir algumas arquiteturas de redes neurais, tais como as redes *perceptron* de múltiplas camadas (MLP, do inglês *multilayer perceptron*), e os modelos recorrentes das LSTMs, GRUs e ESNs. Esse estudo, além de ter um papel fundamental no esquema da investigação associada ao projeto, será de grande importância para a formação de um jovem pesquisador em temas como processamento adaptativo de sinais e inteligência computacional.

Um segundo objetivo deste trabalho é a implementação e a análise sistemática das estruturas neurais no contexto da predição de diferentes tipos de séries - séries com

características sazonais pronunciadas, séries caóticas e séries estocásticas (sinais de informação) – e, em cada caso, serão avaliados aspectos metodológicos desejáveis.

Finalmente, um terceiro objetivo do projeto é a redação de um relatório final que apresente as diferentes estruturas, algoritmos de ajuste de pesos e metodologias de seleção de variáveis de maneira didática e que contenha uma discussão ampla e representativa dos diversos ensaios realizados e dos resultados deles decorrentes. Desejamos produzir um documento que sirva de referência a futuros interessados na área, particularmente no seio do grupo no qual o aluno está inserido.

3 Metodologia

Na primeira etapa do projeto, será realizada uma ampla revisão bibliográfica formada por duas etapas essenciais, a primeira voltada ao estudo de sinais e sistemas [10], processos estocásticos e filtragem adaptativa supervisionada [4], e a outra dedicada às redes neurais selecionadas [11, 12]. Todas as etapas desse estudo serão acompanhadas de implementações práticas pertinentes.

Na segunda etapa, efetuar-se-á um planejamento cuidadoso dos diversos cenários de teste, dos processos de configuração das estruturas de processamento, dos algoritmos de aprendizado adotados e das metodologias de seleção de variáveis [13]. Também serão avaliadas as métricas de análise a serem adotadas de modo a permitir uma adequada comparação dos diversos *setups*. Após esse planejamento, o aluno estará apto a conduzir os testes e a compilar os resultados obtidos, os quais, a seguir, serão alvo de uma criteriosa análise. Ao longo da etapa, haverá sempre uma efetiva preocupação com a validade e a representatividade das análises realizadas, para que a empreitada como um todo conduza a conclusões de valor prático.

Por fim, a terceira e final etapa será devotada ao processo de redação do relatório final, o qual deve apresentar de forma didática as diversas ferramentas empregadas e de maneira sistemática os resultados obtidos e análises empreendidas.

4 Material

Serão utilizados os recursos computacionais do DSPCom (Laboratório de Processamento de Sinais para Comunicações) e do Departamento de Engenharia de Computação e Automação Industrial da Faculdade de Engenharia Elétrica e de Computação da Unicamp. Estarão também disponíveis os recursos associados ao sistema de bibliotecas dessa universidade.

5 Plano de Trabalho

O desenvolvimento deste trabalho se dividirá entre as disciplinas EE015 e EE016.

Para a primeira (EE015), pretende-se realizar a revisão bibliográfica e os estudos teóricos necessários para o entendimento do problema de predição no contexto de séries temporais de sistemas com dinâmica caótica utilizando redes neurais artificiais, incluindo a realização de ensaios com essas ferramentas computacionais de forma a obter uma experiência prática de uso.

Na segunda disciplina (EE016), concluir-se-á a implementação dos métodos de previsão e, então, seus desempenhos serão avaliados comparativamente considerando um conjunto de cenários. Além disso, teremos a elaboração do relatório final, o qual resumirá a trajetória do aluno durante o projeto de iniciação como um todo.

O cronograma a seguir apresenta a divisão e duração das etapas durante o desenvolvimento do trabalho para ambas as disciplinas.

Etapa	Mês					
	1	2	3	4	5	6
I	X	X	X	X	X	X
II					X	X
III						
IV						
V						
VI						

(a) Cronograma de trabalho para EE015

Etapa	Mês					
	7	8	9	10	11	12
I						
II	X					
III	X	X	X			
IV		X	X	X		
V			X	X	X	
VI					X	X

(b) Cronograma de trabalho para EE016

- I. Estudos dirigidos e revisão bibliográfica
- II. Ensaios utilizando redes neurais artificiais para diferentes cenários
- III. Planejamento dos cenários de teste e dos processos computacionais
- IV. Implementação dos modelos preditivos utilizando redes neurais
- V. Estudo comparativo dos diferentes modelos implementados
- VI. Redação do relatório final

A primeira parte do cronograma (etapa (a)) foi concluída e apresentada no relatório da disciplina EE015. Logo, para esta matéria (EE016), serão realizadas e apresentadas no relatório final as atividades indicadas no cronograma (b).

Referências

- [1] J. C. Principe, N. R. Euliano, and W. C. Lefebvre, *Neural and adaptive systems: fundamentals through simulations*, vol. 672. Wiley New York, 2000.
- [2] J. M. T. Romano, R. Attux, C. C. Cavalcante, and R. Suyama, *Unsupervised signal processing: channel equalization and source separation*. CRC Press, 2018.
- [3] D. Erdogmus and J. C. Principe, "From linear adaptive filtering to nonlinear information processing-the design and analysis of information processing systems," *IEEE Signal Processing Magazine*, vol. 23, no. 6, pp. 14–33, 2006.
- [4] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2008.
- [5] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.
- [6] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction," *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 240–254, 1994.
- [7] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "Lstm: A search space odyssey," *IEEE transactions on neural networks and learning systems*, vol. 28, no. 10, pp. 2222–2232, 2016.
- [8] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [9] L. Boccato, *Novas Propostas e Aplicações de Redes Neurais com Estados de Eco*. Tese (doutorado), Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas, SP, 2013.
- [10] A. S. W. w. S. H. Alan V. Oppenheim, *Signals and Systems*. Prentice Hall, 2 ed., 1996.
- [11] S. Haykin, *Neural networks and learning machines*, 3/E. Pearson Education India, 2010.
- [12] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, 2019.
- [13] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.
- [14] R. R. de Faissol Attux, *Novos paradigmas para equalização e identificação de canais baseados em estruturas não-lineares e algoritmos evolutivos*. Tese (doutorado), Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, Campinas, SP, 2005.